BioSci
S
89
.E21
no.521

ATION BULLETIN 521 June, 1982

BIO SCI
LIBRARY

A Methodological Appraisal of the Follow Up Instruments Used in Evaluating Hillsborough County's CETA Programs

by

A. E. Luloff and P. H. Greenwood

NEW HAMPSHIRE
AGRICULTURAL EXPERIMENT STATION
UNIVERSITY OF NEW HAMPSHIRE
DURHAM, NEW HAMPSHIRE

University of New Hampshire
Library

SHOUT
SHARE
IT

BioSci
S
89
.E21
no.521

ATION BULLETIN 521 June, 1982

BIO SCI
LIBRARY

A Methodological Appraisal of the Follow Up Instruments Used in Evaluating Hillsborough County's CETA Programs

by

A. E. Luloff and P. H. Greenwood

NEW HAMPSHIRE
AGRICULTURAL EXPERIMENT STATION
UNIVERSITY OF NEW HAMPSHIRE
DURHAM, NEW HAMPSHIRE

University of New Hampshire
Library

S
89
.E21
no 521

ACKNOWLEDGEMENTS

This publication is a result of the research program of the Institute of Natural and Environmental Resources at the University of New Hampshire. It represents a contribution to Regional Research Project NE-129, "Improving the Distribution of Socio-economic Resources in Rural Areas" and to Project H-266 of the New Hampshire Agricultural Experiment Station.

The authors wish to acknowledge the assistance of George Frick, Tom Ilvento, Judith Wright, Ruth Grady and Ann Stockman in data collection and organization. Special recognition goes to Karen Blandford for her editorial services.

Note: This report is the second in a two-part series. It should be read in conjunction with: "An Evaluation of Economic Gains of Participants in the Hillsborough County CETA Programs," by P. H. Greenwood and A. E. Luloff.

New Hampshire Agricultural Experiment Station Research Report, Number 95, 1982.

Programs of the New Hampshire Agricultural Experiment Station are open to all persons without regard to race, color, national origin or sex. The University of New Hampshire is an Affirmative Action/Equal Opportunity Employer.

UNH LIBRARY
3 4600 00618 4789

ABSTRACT

Follow-up evaluation of several of CETA's major programs during the period January 1, 1979, to September 15, 1980, was conducted by the Institute of Natural and Environmental Resources. An evaluation of the instrument used in the analysis is undertaken to provide guidelines for interpretation of the data in the report. "An Evaluation of Economic Gains of Participants in the Hillsborough County CETA Programs," by P.H. Greenwood and A.E. Luloff. The results indicate that the failure to pretest the instrument contributed to the problems of validity encountered. Further, in the absence of a control group, it is difficult to measure gains or losses by the CETA population. Specific recommendations for conducting future evaluations are offered.

KEY WORDS: CETA, Instrument, Cross-sectional Design, Longitudinal Design Evaluation.

TABLE OF CONTENTS

Page

Introduction 1

Methods and Rationale for Survey Analysis 1

The Instruments: Advantages and Deficiencies 2

The Issue of Experimental Design 4

The Issues of Questionnaire Construction 6

Summary 9

References 10

A Methodological Appraisal of the Follow-Up Instruments: Hillsborough County's CETA Programs

by
A. E. Luloff and P. H. Greenwood[1]

INTRODUCTION

At the finish of the first report "An Evaluation of Economic Gains of Participants in the Hillsborough County CETA Programs," mention was made of shortcomings in the design and rationale of the follow-up procedures. While our focus is on the CETA procedures, the problems we address are generic to survey research. The purpose of this report, then, is to examine the issues of experimental design and questionnaire construction as a basic means for improving mail questionnaire research. Shortcomings of the CETA follow-up instrument are explained and alternative mechanisms for reducing some of the biases associated with these problems are suggested. Other researchers with the same or similar data should profit from the inexpensive procedures suggested herein.

METHODS AND RATIONALE FOR SURVEY ANALYSIS

Pursuant to an agreement with the Hillsborough County Prime Sponsor (SNHS) the Institute of Natural and Environmental Resources, University of New Hampshire (INER) engaged in a follow-up evaluation of the following major programs during the period of January 1, 1979 to September 15, 1980:

Classroom Training	Title II B
Skill Training	Title II B
On Job Training	Title II B
Work Experience	Title II B
Services to Participants	Title II B
Public Service Employment	Title II B, Part D PSE

SNHS agreed to provide the necessary intake information on 275 participants who terminated from these programs during this period. The INER research team attempted to conduct follow-up questionnaires on these terminees commencing on or about their six

[1]Assistant Professor of Community Development and Assistant Professor of Resource Economics, Institute of Natural and Environmental Resources, University of New Hampshire, respectively.

month anniversary date. The information was gathered through the use of standard instruments for follow-up evaluation provided by SNHS.

All information (intake and follow-up) was coded, keypunched, and stored on magnetic tapes. This information was processed on the University of New Hampshire's DEC-10 Computer System, and all analyses were conducted using a standard statistical package (Statistical Package for the Social Sciences — SPSS).

THE INSTRUMENTS: ADVANTAGES AND DEFICIENCIES

The first instrument used in the analysis was the Intake Information Form developed by the Department of Economics at Northeastern University. This 6 page pre-coded instrument provided standard information on all CETA participants which was to be used as the benchmark with which the respondents post-CETA labor market experience could be compared. The following major areas of intake information were included:

1) Demographic-program characteristics, individual social characteristics
2) Pre-CETA Labor Market Experience-summary of 12 month period prior to CETA
3) Most recent job prior to CETA
4) Second most recent job prior to CETA
5) Program and termination information

In sum, a minimum of 120 variables were generated through the Intake Information Form.

Accompanying this form was a Technical Assistance Document provided by SNHS for use in conducting this evaluation. This detailed description of coding procedures was designed to allow for maximum uniformity among New England's prime sponsors intake information. The designers of the instrument and documentation package correctly point out that accurate and complete intake information is critical to the successful operation of the follow-up project.

According to these detailed instructions, the data for the intake forms were to be precoded directly by the Prime Sponsor. This procedure mandates accurate and consistent record-keeping on the part of the prime sponsor. In a situation marked by high rates of turnover among those prime sponsor staff personnel involved in gathering, recording, and storing this information, a significant increase in incorrect data transformations is likely. Further, in the absence of detailed decision-making instructions on the part of the local prime sponsor, similar increases in the collection of incorrect information are likely to occur. While uniform instructions can help to allay some of these problems, they cannot provide answers to all situations.

An additional point deserves special attention. The CETA program is premised on the need to train and/or retrain individuals

for reentry into the labor force with new and/or improved skills. However, the program is not open to everyone. Each individual must meet a minimum standard of criteria for eligibility prior to entry. Whether or not the individuals are cognizant of these criteria impacts directly intake information provided to the prime sponsor. If the subjects have been sensitized to the importance of providing "correct" responses, then serious questions of reliability and validity of data must be raised. One way to help control for this potential bias would be to regather pre-CETA work experience information after termination from the program. These answers might better reflect the realities of their experiences than data gathered at time of entry, where "correct" answers are often viewed as entry criteria. While it is beyond the scope of our current research, we would suggest that a comparative analysis be made of these two entries of data.

The second instrument used in the analysis was the Follow-Up Questionnaire also designed by the Department of Economics at Northeastern University. This 19 page pre-coded instrument was designed to provide detailed information on the past program labor market and educational experiences of the program terminees. In addition, a battery of attitudinal, informational, and knowledge questions are asked which are designed to assess the terminees' personal views of the strengths and weaknesses of the program(s) in which they participated. The questions are arranged into the following 16 sections:

1) Introduction—Demographic Data
2) Pretraining Employment History
3) Supportive Services
4) Respondent's Views of Programs
5) Respondent's Program Experience
6) Summary Post-CETA Information
7) Respondent's First Job After CETA
8) Respondent's Second Job After CETA
9) Respondent's Third Job After CETA
10) Respondent's Fourth Job After CETA
11) Respondent's Current or Last Job
12) For Respondents Who Are Not Currently Employed
13) Respondent's Other Post-Program Experiences
14) For Respondents Who Have Attended School Since Leaving CETA
15) For Respondents Who Have Participated in Another Training Program Since Leaving CETA
16) For Respondents Who Have Served in the Military Since Leaving CETA

In sum, a minimum of 219 variables were generated through the follow-up form.

As with the Intake Information Form, a detailed Technical Assistance Document, provided by SNHS, was made available to INER. This document consisted of both general and specific instructions for carrying out the follow-up interview. While the instructions cover the questions which have the highest probability of leading to errors in data collection, the instrument fails to address issues of design which have a major impact on the reliability and

validity of the entire data set produced. It is to these issues that we now briefly turn.

THE ISSUE OF EXPERIMENTAL DESIGN

In social science research there are basically two designs: (1) cross-sectional, and (2) longitudinal (Zetterberg, 1965). The former is characteristic of most survey research. The latter is represented by the panel or time-series approach. Symbolically, the cross-sectional design would appear as follows:

t_1	t_1
X_1	Y_1
X_2	Y_2
X_3	Y_3
.	.
.	.
.	.
X_n	Y_n

where we measure a sample of n units at time t, with regard to variables X and Y. The longitudinal design, on the other hand, can be depicted as follows:

t_1	t_1	t_2	t_2
X_1	Y_1	X_1'	Y_1'
X_2	Y_2	X_2'	Y_2'
X_3	Y_3	X_3'	Y_3'
.	.	.	.
.	.	.	.
.	.	.	.
X_n	Y_n	X_n'	Y_n'

where we measure a sample of n units at times t_1 and t_2 with respect to X and Y.

The major difference between a cross-sectional and a longitudinal design is the introduction of time as a key factor. A dynamic test of an hypothesis in the cross-sectional framework cannot be done; this research design provides a test of the posited relationship $Y = f(X)$.

The longitudinal design, on the other hand, incorporates and extends this process. Not only can one assess $Y = f(X)$ but one can now compare the "n"units at two points in time, introducing a dynamic aspect to the study ($\Delta Y = \Delta f(X)$). Because it is conceivable that the former test might lead to the acceptance of the original hypothesis, while the latter might lead to its rejection, longitudinal designs are seen as more sensitive than cross-sectional designs.

In order to evaluate the effectiveness of the CETA experience on program enrollees, it would clearly be necessary to use a longitudinal framework. This would allow for the monitoring of key factors thought to influence the degree of success of the program (as

measured by the program's effectiveness in enhancing participants income and employment).

The current procedural framework for conducting the CETA Follow-Up Evaluations measures several variables at 2 points in time. However, it would be extremely difficult to be conclusive about the CETA programs' effectiveness on the basis of the measures. The design now being used most resembles a One-Group Pretest-Post-test Design. This design takes the form: O_1 X O_2, where X is the treatment (in this case, CETA program involvement) and O_1 is a measure of group O at time 1, and O_2 is a measure of group O at time 2. This design is subject to severe threats to internal validity. Internal validity provides answers to the question, "has the experimental stimulus made some significant difference in the ongoing experiment?"

The One-Group Pretest Post-test Design is vulnerable to five major threats to internal validity: (1) history, (2) maturation, (3) testing, (4) instrument decay, and (5) statistical regression. History refers to the unique events which occur between t_1 and t_2 other than the stimulus. Maturation refers to the biological or psychological processes which systematically vary over time independent of the stimulus or other external event. Testing refers to the effects attributable to the pretest itself. Whenever the testing process becomes a stimulus for change rather than a barometer of some trait, a reactive effect occurs. Instrument decay or instrumentation refers to the autonomous changes in the measuring instrument which might account for the changes on O_1 - O_2. These processes include learning (recall) and fatigue (when recording responses) among respondents. Statistical regression occurs whenever experimental groups are selected on the basis of uniqueness or extremity. Extreme scores, in part, are seen to reflect random errors in sampling techniques and thus promote random instability of measurement. Regression effects are often inevitable accompaniments of imperfect test-retest correlation for groups selected for their extremity (Campbell and Stanley, 1963:11).

Because of these design problems, internal validity is severely limited. Without internal validity, measures of external validity become less reliable. External validity is concerned with which populations, settings, and variables the effect(s) under study can be generalized. In scientific research we will most commonly sacrifice external validity for internal validity. Unfortunately, the present CETA Follow-Up Evaluation Design is weak in both areas.

In order to help address some of these problems we would suggest the use of a different and logically superior design commonly referred to as the Solomon Four-Group Design. This design explicitly takes into consideration both external and internal validity factors. The structure of the design is as follows:

R	O_1	X	O_2
R	O_3		O_4
R		X	O_5
R			O_6

where R stands for randomization (which is employed to achieve equivalent groups and is a direct attempt to control for the above mentioned internal validity problems); X is the treatment (CETA program experience); and O_n represents the groups for comparative purposes. This design examines external validity in that the introduction of groups O_5 and O_6 (no pretest) allow for comparisons between groups O_1 to O_4 with respect to the main effects of testing and the interaction of testing and X (CETA program experience). Its major advantage is that it allows for a control group (O_3, O_4) to compare with the treatment group. In fact, the first two lines represent the classic experimental design, a luxury not common to most social science research.

Upon examination of the structure of the design, four different comparisons of the effect of X are possible: (1) $O_2 > O_1$; (2) $O_2 > O_4$; (3) $O_5 > O_6$; and (4) $O_5 > O_3$. According to Campbell and Stanley (1963: 25):

> The actual instabilities of experimentation are such that if these comparisons are in agreement, the strength of inference is greatly increased.

In the analysis of the effects of this design, a 2 x 2 analysis of variance is possible:

	NO X	X
Pretest	O_4	O_2
No Pretest	O_6	O_5

To perform the statistical test one would estimate the main effects of X from the column means, the main effect of pretesting from the row means, and the interaction of testing with X from the cell means.

The differences between the two designs (one-Group Pretest-Post-test vs. Solomon Four-Group Design) are apparent. Control is a key element of the experimental designs whereas in the One Group Pretest-Post-test Design such control is lacking. Causality, in its true sense, is difficult to achieve at any time. However, if we are interested in speaking to this question, the One-Group Pretest-Post-test Design is clearly not applicable. With the Solomon Four-Group Design, the comparative measures necessary to overcome the internal validity threats characteristic of the One Group Pretest-Post-test Design are available.

THE ISSUES OF QUESTIONNAIRE CONSTRUCTION

While experimental design issues are a major problem in the Follow-Up Evaluation Project, they are not the only pressing issues which need to be addressed. Of primary importance among these latter issues are problems which relate to the construction of the actual instruments (questionnnaires). Some of the most important issues which will be discussed include tracking, length,

number of contacts made, sponsorship of instrument, and salience.

Tracking or locating respondents from an earlier time period for the purposes of gathering longitudinal data on them is a difficult and often time-consuming experience. In the previous report, "An Evaluation of Economic Gains of Participants in the Hillsborough County CETA Programs," intake information was compiled by SNHS which included respondents current addresses and phone numbers. A minimum of six months elapsed prior to the onset of the follow–up questionnaire procedures. If the respondent changed his/her address or phone number and the change was not recorded in the intake data file, then, obviously, tracking the respondent became more difficult. To overcome this problem we called the original number seeking information on new phone numbers or addresses where the respondent could be reached. In addition, "address correction requested" envelopes were mailed to each respondent's former address so that the Post Office could supply us with new forwarding addresses when possible. However, both procedures resulted in long time delays and costly labor overuns. We would suggest that periodic updating of intake information (address and phone number) be instituted at SNHS to help overcome some of these problems. Further, we would suggest that on all new intake forms a new section be added which asks for the phone number and address of the respondents nearest relative or friend who is an established resident (at least 5 years) of New Hampshire. This latter procedure should help to insure the gathering of complete information on the respondents current location.

Prior to conducting the actual survey, SNHS agreed to notify all terminees through the mail, of the survey efforts. This letter, signed by James Machakos, CETA Administrator for SNSH, informed the terminees of the follow–up evaluation and requested their co-operation. This initial contact served two main purposes: (1) the endorsement of a major executive administrator of the program lent credibility to the survey effort, and (2) by encouraging the terminees to participate in the follow up effort, the level of salience among the potential respondents was heightened. The letter was mailed approximately two weeks prior to the administration of the actual follow-up survey instrument. Unfortunately this procedure did not greatly enhance the overall completion rate (aproximately 48%).

The low completion rate suggests many potential hazards in interpretation of this data. Some of these potential problems revolve around the instrument itself. Of most importance is the problem of question and questionnaire length. Questions which are too long often contain several ideas which tend to cause misunderstanding between the respondent and interviewer. This often results from the attempt to transfer questions used in a person-to-person or mail survey to a telephone survey. We suggest the shortening and simplification of many of the questions. An additional aid could be accomplished by offering a "key word summary" which helps to focus the questions while putting them into a more conversational form.

The problem of length is not limited to the questions, however. Often, as in the CETA Follow-Up Instrument, there is a potential need for identifying multiple gradations among response categories. When several response categories are used, two major problems are often encountered: (1) remembering the categories (reminding respondent of them can help here), and (2) the difficulty of finding the proper words to connote intensity of feelings of response. The latter problem is best addressed by sacrificing some of the information by reducing the number of categories. This problem and its potential solution is most prevalent among open ended, attitudinal, opinion type, and historical recall questions (most prominently in this survey the questions dealing with labor history). Question ordering is also very important. Not only should similar and related questions be grouped together into sections, but these sections should be logically related so as to provide for a smooth transition between and among them. Thus, the context of the instrument plays an important, albeit unobtrusive, role in conveying to the respondent that this procedure is worthwhile. In the CETA Follow-Up Evaluation used in the Greenwood-Luloff report, this problem is critical. For example, since the measurement of any improvements in employment history among respondents following CETA participation is a main goal of the survey, it would appear that a question which directly asks for the duration of the respondents employment (or unemployment) during said period would appear critical.

However, in order to arrive at this(ese), figure(s), it was first necessary to go through a 6-step procedure. The lack of specific questions, especially in the presence of detailed minutia, contributes to the confusion indicated by the respondents, coders and researchers in this study. It appears that too much time was spent on seeking specific answers without regard for the overall or "big picture."

In addition, it should be pointed out that the first few questions of an instrument are the most critical. There are basically three reasons for this: (1) if a response can be obtained to these questions, the likelihood of a subsequent termination prior to completion of the instrument is greatly reduced, (2) these questions can be used to divert objections to the survey by drawing the respondents attention away from other concerns, and (3) the way these questions are asked and answered tends to set the tone for the ensuing interview (Dillman, 1978:218). Therefore, it is necessary to make these first few questions interesting and relevant to the respondents. This is not accomplished by beginning with demographic or background questions (which should be the last ones asked in a telephone interview) as in the CETA Follow-Up Survey.

When detailed information is necessary to track a respondents history with respect to certain variables, a careful and straight forward design is needed. The use of "screening questions" might help here. All questions which refer to all respondents should begin

on the left margin of the instrument. As the screening process begins, i.e., as you encounter questions which refer only to a segment of the respondents, indentations (normally 5 to 10 spaces) from the margin are used. This alerts the coder, keypuncher, and researcher to the fact that only a segment of the respondents are providing information to these questions. Such a procedure allows for a smoother transition from section to section. It also helps to alleviate some of the problems which normally are encountered when asking recall or historical information by providing the interviewer with handy "prompts" or reminders should the respondent indicate confusion over the question. Such a procedure would be of invaluable assistance in the current CETA Follow-Up Instrument.

One final suggestion for future surveys is offered. We believe that all instruments (mail, telephone, and/or person-to-person questionnaires and schedules) should be thoroughly pretested prior to the onset of the survey instrument. It is a means of assessing how the questionnaire works and whether any changes are necessary prior to the start of the full-scale study. A good pretest is designed to test not only the questionnaire, but the questions as well. If such a pretest had been conducted prior to the SNHS sponsored research, we believe many of the problems herein reported could have been avoided.

SUMMARY

This report addresses common methodological problems facing survey researchers. The design of a valid and reliable instrument for the collection of social science data is a necessary first step to survey research. Unfortunately, not enough attention is usually given to issues of experimental design.

The use of a control group for comparisons with an experimental or treatment group greatly enhances the utility of a survey instrument. In the usual case, that is, a One-Group Pretest-Post-Test Design, several major threats to internal validity arise. These include the effects of history, maturation, testing, instrument decay, and statistical regression. By introducing a control group and using a randomization procedure for the assignment of study groups, these threats are controlled. Further, the introduction of the control group allows for a more meaningful analysis of relationships.

Additionally, the survey scientist needs to focus his attention on issues which relate directly to questionnaire construction. A procedure for locating respondents must be developed prior to the start of a longitudinal survey. Endorsements from relevant groups should be gathered prior to the collection of data to help insure a high response rate. The instrument itself needs to be worded so as to not confuse the respondents, especially when the data gathering

techniques are either mail or telephone. Questions need to be ordered so as to encourage responses from the targeted audiences. Starting a questionnaire with such things as age, education, home town, and marital status may be appropriate in some questionnaires, but not in others (as was the case in the CETA instrument). The first questions should be relevant and interesting to the respondents. Finally, we suggest that all survey researchers pretest their instruments. This procedure helps to insure that the instrument can be understood by the respondents and that it gathers the kind of information it was designed to collect.

REFERENCES

Campbell, Donald T. and Stanley, Julian C., 1963. *Experimental and Quasi-experimental Designs for Research.* Rand McNally and Company, Chicago, Ill.

Dillman, Don A., 1978. *Mail and Telephone Surveys.* John Wiley and Son, New York.

Greenwood, P.H. and Luloff, A.E., 1982. "An Evaluation of Economic Gains of Participants in the Hillsborough County CETA Programs." New Hampshire Agricultural Experiment Station, Research Report No. 95.

Nie, Norman, Hull, C. Hadlai, Jenkins, Jean G., Steinbrenner, Karin and Bent, Dale H., 1975. *Statistical Package for the Social Sciences.* McGraw-Hill, New York.

Zetterberg, Hans L., 1965. *On Theory and Verification in Sociology,* third edition. Bedminster Press, Totwa, New Jersey.